Subsemigroups of a free semigroup and the graph lemma

While the subgroups of a free group are themselves free, the subsemigroups of a free semigroups need not be free. For example, the semigroup $S = \langle X \rangle$ generated by the set $X = \{a, ab, ba\}$is not free since $ab \cdot a = a \cdot ba$, and X is the smallest set generating S. In general, we have the following fact:

Lemma. Let S be a subsemigroup of a free semigroup. Then S has a smallest (w.r.t. inclusion) set of generators $B = S \setminus S^2$.

Proof. We first show that B generates S. Assume to the contrary that $w \in S$ but $w \notin \langle B \rangle$, and let w be such a word of minimal possible length. Since $w \notin S \setminus S^2$, we have $w \in S^2$, that is, $w = uv$ for some $u, v \in S$. The minimality of $|w|$ implies that $u, v \in \langle B \rangle$, therefore also $w \in \langle B \rangle$, a contradiction.

Let now $S = \langle B' \rangle$. In particular $B \subseteq \langle B' \rangle$. The definition of B implies, that no element of B is a product of two or more elements of B'. Hence $B \subseteq B'$. □

The set B of the previous lemma is called the *basis* of the semigroup S, and its size is the *rank* of S. The subset T of the semigroup $\{a, b\}^*$ consisting of words starting with a shows that a semigroup of finite rank can be of an infinite rank. Namely, the basis of T is the set $\{ab^i \mid i \geq 0\}$.

Let us stress again that while the basis is the smallest generating set of a given semigroup, the semigroup need not be free. If B generates a free semigroup, then it is called a *code*.

Next lemma characterizes semigroups generated by a code.

Lemma. A semigroup $S \subseteq \Sigma^+$ is free iff for any $p, q, w \in \Sigma^+$ we have

$$p, q, pw, wq \in S \Longrightarrow w \in S\,. \tag{f}$$

Proof. Let S be free and let $p, q, pw, wq \in S$. Then also $pwq \in S$ and the words pw, wq a pwq have a unique factorization into elements of the basis B_S of S. Let $p = p_1 p_2 \cdots p_{i_p}$, $q = q_1 q_2 \cdots q_{i_q}$, $pw = b_1 b_2 \cdots b_{j_1}$ and $wq = c_1 c_2 \cdots c_{j_2}$ be such factorizations (that is, all p_i, q_i, b_i and c_i are from B_S). Then the equality

$$p_1 p_2 \cdots p_{i_p} c_1 c_2 \cdots c_{j_2} = pwq = b_1 b_2 \cdots b_{j_1} q_1 q_2 \cdots q_{i_q}$$

implies $p_k = b_k$, $k = 1, 2, \ldots, i_p$, hence $w = b_{i_p+1} b_{i_p+2} \cdots b_{j_1} \in S$.

Let now S be not free and let $b_1 b_2 \cdots b_j = c_1 c_2 \cdots c_k$ is a shortest possible nontrivial relation between elements of B_S. WLOG, let $b_1 < c_1$. Then $p = b_1$, $q = c_2 c_3 \cdots c_k$ and $w = b_1^{-1} c_1$ do not satisfy (f). □

The implication (f) is called the *stability condition*.

Since sets satisfying th stability condition are clearly closed under the intersection, there is a smallest (w.r.t. inclusion) free semigroup F containing a given set X. Such a semigroup is called the *free hull* of the set X, and we write $F = \langle X \rangle_{\mathtt{f}}$. The basis F is called the *free basis* of the set X and its cardinality, denoted $\mathrm{rank}_{\mathtt{f}}(X)$, is called the *free rank* of the set X.

Note that the stability condition can be written as

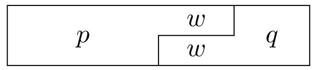
$$wS \cap S \neq \emptyset \quad \& \quad Sw \cap S \neq \emptyset \quad \Longrightarrow \quad w \in S\,.$$

This is equivalent ot a seemingly stronger

$$wS \cap S \cap Sw \neq \emptyset \quad \Longrightarrow \quad w \in S\,,$$

since $wpqw \in wS \cap S \cap Sw$ if $p, q, pw, wq \in S$.

Note also the graphical meaning of the stability condition. It says that there can be no nontrivial relation by requiring that any "overflow" in a relation be included into the free hull:



We now have an algorithm for obtaining the free basis of a finite set X: Let $b_1b_2\cdots b_j = c_1c_2\cdots c_k$ be a nontrivial relation of elements from X. If the relation is minimal (not composed of shorter relations), then we can assume, by symmetry, that $|b_1| < |c_1|$. In X, replace c_1 with $c_1' = b_1^{-1}c_1$. By the stability condition, the new set X' has the same free hull as X. The process terminates by induction on the total length of X with the free basis B_X.

It is clear from this algorithm that the free rank of X is at most $|X|$. Moreover, the free rank is strictly less if X is not itself a code. This follows from the fact that replacing each c_1 with b_1c_1', in the nontrivial relation above, yields a nontrivial relation again unless the original relation was $b_1c_1' = c_1$. In such case, however, we just remove c_1, hence $|X'| < |X|$.

The following lemma turns out to be very useful.

Lemma. Let X be a set of words, and let B be the free basis of X. Then for each $b \in B$ there is $x \in X$ such that b is the first (the last resp.) factor of the B-factorization of x.

Proof. If X is finite, the claim follows easily by induction from the above algorithm: 1. It is trivial for X being a code. 2. If X' has the property, then also X has it since the first B-factor of c_1' is also the first B-factor of b_2.

For a possibly infinite X, the proof goes by contradiction. Assume that $b \in B$ is not the first factor of the B-factorization of any $x \in X$. Let

$$Z = (B \setminus \{b\})b^* = \{cb^i \mid b \neq c \in B\}.$$

Then Z is a code, since the unique B-factorization of each $w \in \langle Z\rangle$ yields a unique Z-factorization of w. Since $X \subseteq \langle Z\rangle \subsetneq \langle B\rangle$, we have a contradiction with the minimality of $\langle B\rangle$. $\square$

We now easily obtain an important theorem called the "Defect theorem" or the "Graph lemma".

Theorem. Let the words from $X = \{w_1, w_2, \ldots, w_n\}$ satisfy relations $(u_i, v_i) \in \Xi^+ \times \Xi^+$, $i \in I$, where $\Xi = \{x_1, \ldots, x_n\}$. Let $G = (X, H)$ be an undirected graph with edges

$$H = \{\{\mathrm{pref}_1(u_i), \mathrm{pref}_1(v_i)\} \mid i \in I\}.$$

Then $\mathrm{rank}_{\mathtt{f}}(X)$ is at most the number of connected components of G.

In particular, if X is not a code, then $\mathrm{rank}_{\mathtt{f}}(X) < |X|$.

Proof. Let B be the free basis of X and let b_i be the first B-factor of u_i. By the previous lemma, we have $B = \{b_1, b_2, \ldots, b_n\}$.

Let $\psi : \Xi^+ \to X^+$ be the morphism defined by $\psi(x_i) = w_i$. Let $\{x_j, x_k\} \in H$, and let $x_j = \mathrm{pref}_1(u_i)$ and $x_k = \mathrm{pref}_1(v_i)$. Since the word $\psi(u_i) = \psi(v_i)$ has a unique B-factorization, we have $b_j = b_k$. The claim follows. $\square$